

FIGURE 1

The Universal Code Equivalent of
the Mitochondrial 1-Sce 1 Gene.

AAAAAATAAAATCAT ATG AAA AAT ATT AAA AAA AAT CAA GTA ATG AAT CTC GGT GGT AAT TCT
 M K N I K K N Q V M N L S P N S
 AAA TTA TTA AAA GAA TAT AAA TCA CAA TTA ATT GAA TTA AAT ATT GAA CAA TTT GAA GAA
 K L L K E Y K S Q L I E L N I E Q F E A
 GGT ATT GGT TTA ATT TTA GGA GAT GCT TAT ATT CQT AGT CQT GAT GAA GGT AAA ACT TAY
 G I G L I L G D A Y I R S R D E S K T Y
 TGT ATG CAA TTT CAC TCC AAA AAT AAG GCA TAC ATG GAT CAT GTA TGT TTA TTA TAT GAT
 C M Q F E W K N K A Y M D M V C L L Y S
 CAA TGG GTA TTA TCA CCT CCT CAT AAA AAA GAA AGA GTT AAT CAT TTA GGT AAT TTA GTA
 Q W V L S P P H K K E R V N H L Q N L V
 ATT ACC TGG GGA GCT CAA ACT TTT AAA CAT CAA GCT TTT AAT AAA TTA GCT AAC TTA TTT
 I T W G A Q T F K H Q A F N K L A N L F
 ATT GTA AAT AAT AAA AAA CTT ATT CCT AAT AAT TTA GTT GAA AAT TAT TTA ACA CCT ATG
 I V N N K K L I P N N L V E N Y L T P H
 AGT CTG GCA TAT TGG TTT ATG GAT GAT GGA GGT AAA TGG GAT TAT AAT AAA AAT TCT CTT
 S L A Y W F H D D Q G K W D Y N K N S L
 AAT AAA AGT ATT GTA TTA AAT ACA CAA AGT TTT ACT TTT CAA GAA GTA GAA TAT TTA CTT
 N K S I V L N T Q G F T F E E V C Y L V
 AAA GGT TTA AGA AAT AAA TTT CAA TTA AAT TGT TAT GTT AAA ATT AAT AAA AAT AAA CCA
 K G L R N K F Q L N C Y V K I N K N K P
 ATT ATT TAT ATT GAT TCT ATG AGT TAT CTG ATT TTT TAT AAT TTA ATT AAA CCT TAT TTA
 I I Y I D S H S Y L I F Y N I I K P Y L
 ATT CCT CAA ATG ATG TAT AAA CTG CCT AAT ACT ATT TCA TCC GAA ACT TTT TTA AAA TAA
 I P Q M M Y K L P N T I S S E T F L K

00492697-012700

FIGURE 2

The synthetic I-Sce I gene

Bam HI

CCGGATCCATG CAT ATG AAA AAC ATC AAA AAA AAC CAG GTA ATG AAC CTG GGT CCG AAC TCT
H H M K N I K K N Q V M N L G P N S
AAA CTG CTG AAA GAA TAC AAA TCC CAG CTG ATC GAA CTG AAC ATC GAA CAG TTC GAA GCA
K L L K E Y K S Q L I E L N I E C F E A
1. GGT ATC GGT CTG ATC CTG GGT GAT GCT TAC ATC CQT TCT CGT GAT GAA GGT AAA ACC TAC
G I G L I L G D A Y I R S R D E G K T Y
TGT ATG CAG TTC GAG TCG AAA AAC AAA GCA TAC ATG GAC CAC GTA TGT CTG CTG TAC GAT
C M O F E W K N K A Y M D H V C L L Y D
CAG TCG GTA CTG TCC CCG CCG CAG AAA AAA GAA CQT GTT AAC CAC CTG GGT AAC CTG GTA
Q W V L B P P H K K E R V N H L B N L V
ATC ACC TGG GGC GCC CAG ACT TTC AAA CAC CAA GCT TTC AAC AAA CTG GCT AAC CTG TTC
I T W G A Q T F K H Q A F N K L A N L F
ATC GTT AAC AAC AAA AAA ACC ATC CCG AAC AAC CTG GTT GAA AAC TAC CTG ACC CCG ATG
I V N N K K T I P N N L V E N Y L T P M
2. TCT CTG GCA TAC TGG TTC ATG GAT GAT GGT GGT AAA TGG GAT TAC AAC AAA AAC TCT ACC
S L A Y W F M O O G G K W O Y N K N S T
AAC AAA TCG ATC GTA CTG AAC ACC CAG TCT TTC ACT TTC GAA GAA GTA GAA TAC CTG GTT
N K S I V L N T Q S F T F E E V E Y L V
AAC GGT CTG CQT AAC AAA TTC CAA CTG AAC TGT TAC GTA AAA ATC AAC AAA AAC AAA CCG
K G L R N K F Q L N C Y Y K I M K N K P
ATC ATC TAC ATC GAT TCT ATG TCT TAC CTG ATC TTC TAC AAC CTG ATC AAA CCG TAC CTG
I I Y I D S M S Y L I F Y N L I K P Y L
ATC CCG CAG ATG ATG TAC AAA CTG CCG AAC ACT ATC TCC TCC GAA ACT TIC CTG AAA TAA
I P Q M M Y K L P N T I S S E T F L K
TAAGTCGACTGCAGGATCCGGTAAGTAAGTAA
Sal I Pst I Bam HI

1 and 2: These amino acids are absolutely necessary to produce catalytic activity. Other substitutions are possible, such as deletions of the 10 first amino acids.

004270-6926460

FIGURE 3

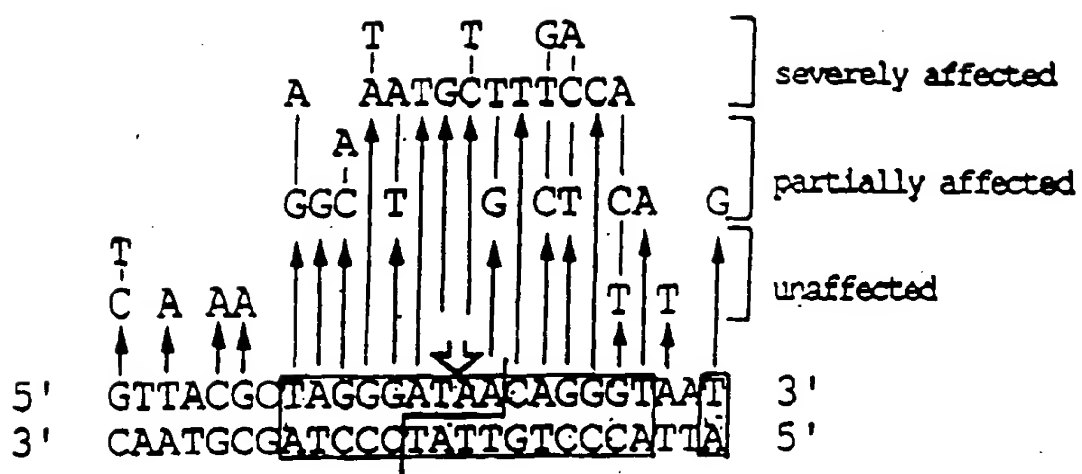


FIGURE 4

1667 CCGACACCTGCTTATTTCTGACGCTGCAACAGCAAGGCAAGGAGCTTCAAGGCAAGGCTTCTACCT 1674
1747 TTATATCTCTCTGCGGTTTCTGACGCTGCAACAGGCTGCAATTTCTG AAG CTC CTC AAG GAG GAG GAG 1818
1819 CTT ATG GAA AAA CGC CAG CAA CGC GCG CTT TTT ACT GTT CCG GCG CTT TCG CTC GCG CTT 1874
1879 TCA TCA CAT GTT CTT TCG GTT ATC CCG TCA TCTGTGATACGCTATTTCTGCTTCTGACG 1947
1948 TGAACCTGCTTCAACGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAG 2017
2020 CCGCTCTCGCGCGGCTTCTGATTTATTA ATG CAG CTC GCA CCA CAG GTT TCG GCA CTC GAA AGC 2084
2088 CCG CAG TCA GCGAAGGCAATTA ATG TCA GTTATCTACTTATACGCTGCAAGGCTTCTACCTT ATG 2164
2145 CTT CCG CCG CCG ATC TCG TCT GGA ATT GTS AGC GGA TAA GAATTTACACAGGCAAGGCT ATG 2228
2229 ACC ATG AAT AAT AAT TCT CAT CTT TCA CAGCTTATCTGATACGCTTAA ATG GAG TCTATCAC 2295
2296 AGTTAAATCTTACGCAAGGCAAGGCAAGGCT ATG AAA TCT AAC AAT GCG CTC ATC CTC ATC CTC GCG 2363
2364 ACC CTC ACC CTC CAT GGT GTA GCG AAT GCG TCG GGT AAG CCG GTA CTC CCG CCG CTC TCG 2423
2424 CCG GAT AAG CCG CTC ATG CTT GAA CCG GAC GGA CCG AAC CAG CCG GAC AAG TCT CTC CTC 2483
2484 TTU GAT TCG GCA TCG CAG CAC AAC TTC TCG TCG GGT AAC CTC CTC ACC CCG GCG AAG CTT 2543
2544 ACT CCG CAT GGU CTT GTT CAC AAT TAA TCACTGCTGCTATTA ATG TGT GCA ATT CTC ACC GGA 2606
2607 TAA CAATTCACAGGCAAGGCAAGGCT ATG CAT ATG AAA AAC ATC AAA AAA AAC CAG GTA ATG 2670
2671 AAC CTC CCG CCG AAC TCT AAA CTA CTA AAA GAA TAC AAA TCG CAG CTC ATC CAA CTC AAC 2730
2731 ATC GAA CAG TTT GAA GCA CCG ATC CCG CTC ATC CTA GGT GAT GGT TAC ATC GGT TTT CTT 2790
2791 GAT GAA CCG AAA ACC TAC TCT ATG CAG TTC CAG TCG AAA AAC AAA GCA TAC ATC GAC CAC 2850
2851 GTA TGT CTC CTC TAC CAT CAG TCG GGT CTC TCG CCG CCG CAG AAA AAA GAA CCG GTT AAC 2910
2911 CAG CTC CCG AAC CTC GGA ATC ACC TCG CCG CCG CAG ACT TTC AAA CAG CAA CCG TTC AAC 2970
2971 AAA CTA CCG AAC CTC TCG ATC GTT AAC AAC AAA AAC ACC ATC CCG AAC AAC CTC GTT CAA 3030
3031 AAC TAC GGT ACC CCG ATG TCT CTC GCA TAC TCG TTC ATG GAT GAT GGT GGT AAA TCG CAT 3090
3091 TAC AAC AAA AAC TAC AAC AAC AAA TCG ATC GTA CTC AAC ACC CAG TCT TTC ACT TTC CAA 3150
3151 CAA GGA GGA TAC CTC GGT AAA GGT CTA GGT AAC AAA TTT CAA CTC AAC TGT TAC GTA AAA 3210
3211 ATC AAC AAA AAC AAA CCG ATC ATC TAC ATC GAT TCT ATG TCT TAC GTC ATC TTC TAC AAC 3270
3271 CTC AAC AAA CCG TAC CTC ATC CCG CAG ATG ATG TAC AAA CTA CCG AAC ACT ATC TTT TTT 3330
3331 GAA ACT TTC CTC AAA TAA TAACTGCACTGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAG 3404
3404 GAA ACT TTC CTC AAA TAA TAACTGCACTGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAG

I-SceI coding sequence of pSCM525 - Note the two amino acid N-terminal extension as compared to genuine version of the gene.

FIGURE 5

VARIATIONS AROUND THE 1-Scp 1 SEQUENCE

	-2	-1	1				5					10							
	H	H	M	K	N	I	K	K	N	I	V	N	N	L	G	P	V	S	
			20									30							
K	L	L	K	E	Y	K	S	D	L	I	E	L	N	I	E	S	F	E	A
		40										50							
G	L	G	L	L	L	L	A	Y	L	R	S	R	D	E	S	K	T	T	
		60										70							
C	M	Q	F	E	W	K	N	K	A	Y	M	D	H	V	C	L	-	V	D
		80										90							
Q	W	Y	L	S	P	P	H	K	K	E	R	V	N	H	L	G	N	-	V
		100										110							
I	T	W	G	A	D	T	F	K	H	Q	A	F	N	K	L	A	N	-	F
		120										130							
I	V	N	N	K	K	I	I	P	N	N	L	V	E	N	Y	L	T	P	-
		140										150							
D	L	A	Y	W	F	H	D	D	G	G	K	W	D	Y	N	K	N	S	Z
		160										170							
N	K	S	I	V	L	N	T	Q	S	F	T	F	E	E	V	E	Y	-	V
		180										190							
K	G	L	R	N	K	F	Q	L	N	C	Y	V	K	I	N	K	N	K	P
		200										210							
I	I	Y	I	D	G	M	S	Y	L	I	F	Y	N	L	I	K	P	Y	-
		220										230							
I	P	Q	H	N	Y	K	L	P	N	T	I	S	S	E	T	F	L	K	-

Positions that can be changed without affecting enzyme activity (demonstrated)
 positions -1 and -3 are not natural. The two amino acids are added due to cloning strategies
 positions 1 to 10: can be deleted
 position 36: G is tolerated
 position 40: M or Y are tolerated
 position 41: S or N are tolerated
 position 43: A is tolerated
 position 46: Y or N are tolerated
 position 91: A is tolerated
 positions 133 and 158: L are tolerated
 position 222: A and S are tolerated

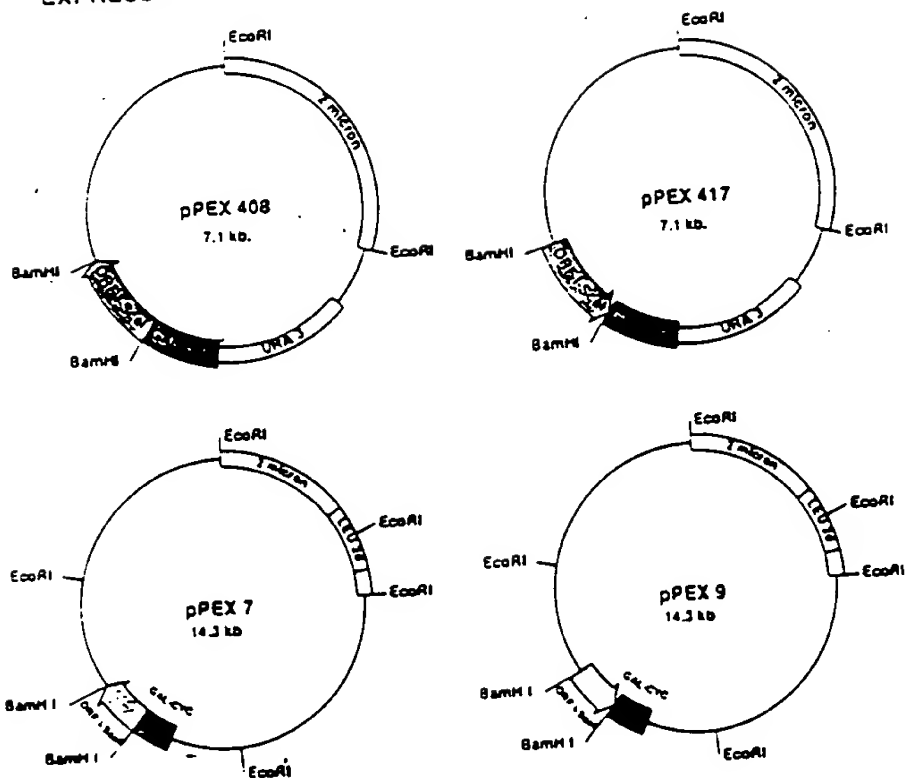
Changes that affect enzyme activity (demonstrated)
 position 19: L to S
 position 38: I to S or N
 position 39: G to D or R
 position 40: L to R
 position 42: L to R
 position 44: D to E, Q or H
 position 45: A to E or D
 position 46: Y to D
 position 47: I to R or N
 position 50: L to S
 position 144: D to E
 position 145: D to E
 position 146: Q to E

00492697-012700

Endonuclease	Recognition Sequence	Cleavage Site	Intron site
I-Sce I (<i>Saccharomyces mitochondria</i>)	CGC TAGGGATAA CAG GGTAA TATAGC GCC ATCCC TATTTTCG AATATAGC		
I-Sce IV (<i>Saccharomyces mitochondria</i>)	TTCTCATGATTCAGCTCTAATCCATG AAGAGTACTTAATCAGATAGGATAGC		
I-Sce II (<i>Saccharomyces mitochondria</i>)	CTTTGGTCAATCAGAAAGTA TATATTT GAAAGCAGTAGGTTTCTATATATAT		
I-Ceu I (<i>Chlamydomonas chloroplast</i>)	TAACGGTCCCAA GTTAGCCAAATTA ATTGCCAGGATTCATCGCCTTAAT		
I-Hpo I (<i>Physarum nucleus</i>)	TGACTCTCTTAAAGGATAGCC AAATGCC ACTGAGAGCAATCCCATCTGT TATAT		
I-Sce III (<i>Saccharomyces mitochondria</i>)	CGAGGTTTTGTAACTATTTACTAGC CCTCCRAAAACUATTUATAAATAATGG		
I-Cre I (<i>Chlamydomonas chloroplast</i>)	GGGTTCAAAACGTCTGTAGACAGTTT CCCAAGTTTTTCAGACACCTCTGCCAAA		
Endo Scc I (RF3) (<i>Saccharomyces mitochondria</i>) (Non isozyme)	GATGCTGTAGGC ATAGGCTTGGTTAT GTACGACATCCGTATCCGAACCAATA		
HO (<i>Saccharomyces nucleus</i>) (Non isozyme)	CTTTCGGCAACAGTATATATTTTATAA CAAAAGCCCTCTTATATTAATAATAT		
I-Cam I (<i>Chlamydomonas mitochondria</i>) (Positive endonuclease)	ACCATGGGGTCAAATGTTTCTTCTGGG TGGTACCCCAATTTACAGAAAGACCC		
I-Pan I (<i>Podosporea mitochondria</i>) (Partial or endonuclease)	GTGCTGAAACATATTTTATTTCTCTT CAGCCAGCTACTATTAATAAAGGAAA		
(<i>Bacteriophage T4</i>)			
I-Tev I	CAACGCCGCAATGATTTTTCTTGGTCTACCGTTTAA GTTGCCAATGATCTTAATAAAGAACCCAGATGGCAATTA		
I-Tev II	CAAGCTTATGATATGAAGTGAACACGTTTAT GTTGAAATACTCATACTTCACCTTGTGCAATTA		
I-Tev III	GCTATTCGTTTCTATGATATCTTCTGCTGACCTTTAA CGATAGCAATAATCATAGAAACCCACATGCAAAAT		

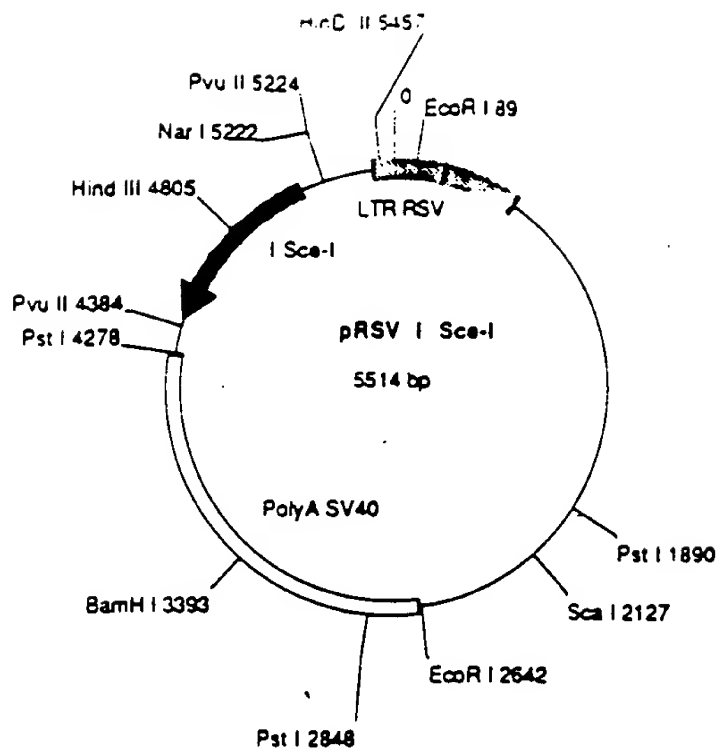
FIGURE 7

EXPRESSION VECTORS



00492697-012700

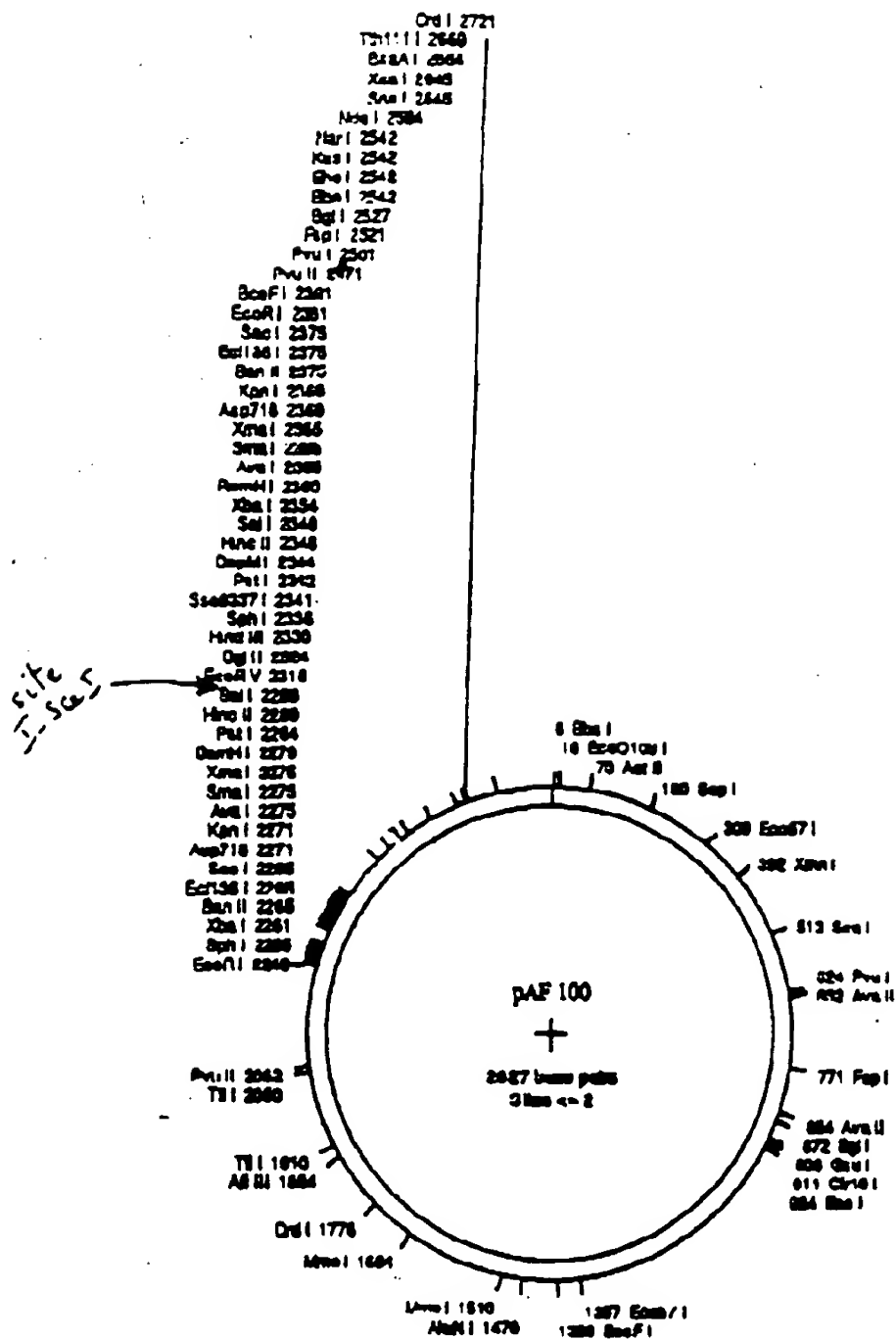
FIGURE 8



002270-26926460

FIGURE 9

0049697.012700



007070-010

[illegible]

PAF 100 - RESTRICTION MAP

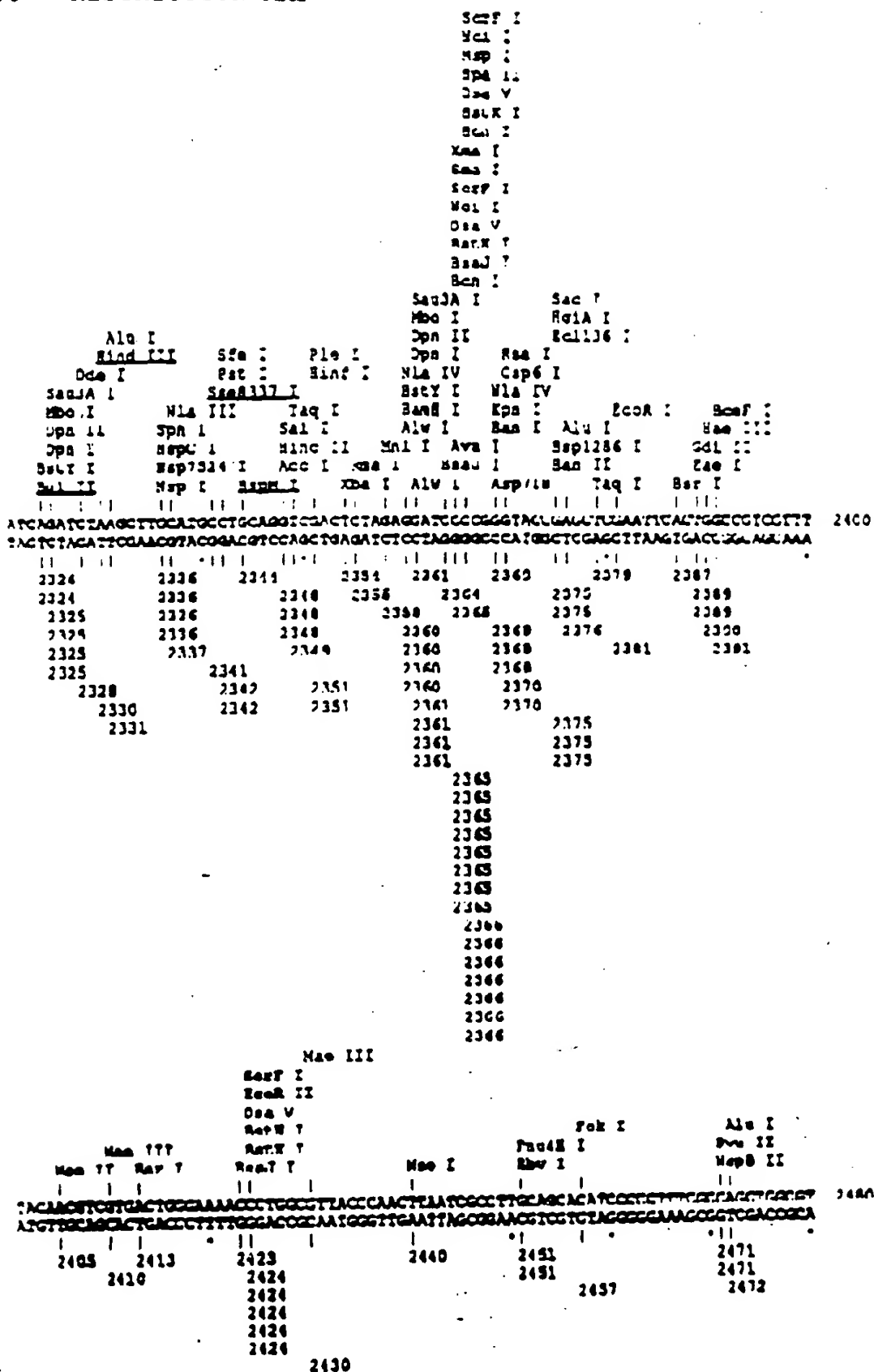
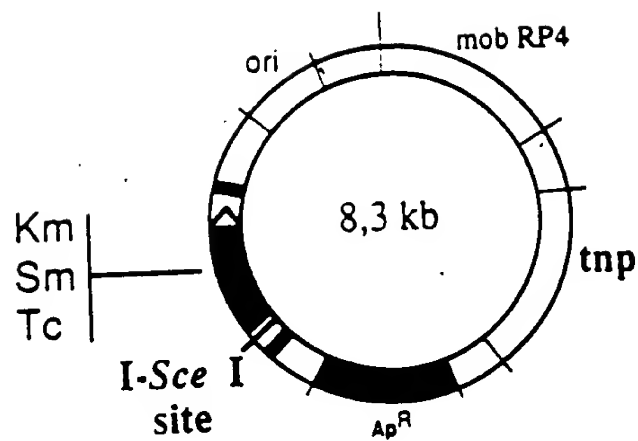


FIGURE 11

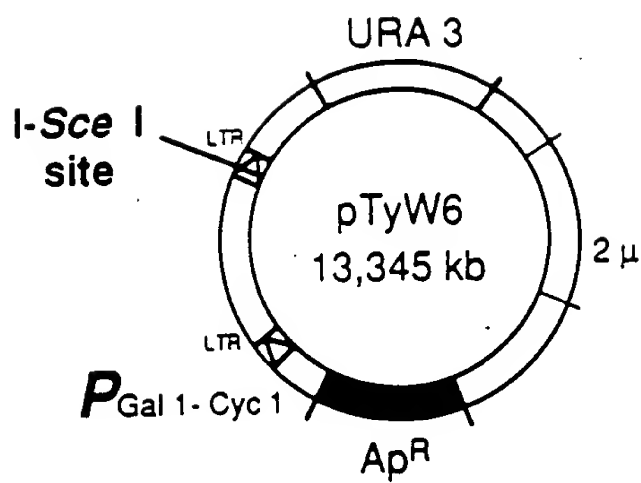
Names : pTSm ω
pTKm ω
pTTc ω



Construction : pGP 704 from De Lorenzo, with transposase gene and insertion of the linker [I-SceI] in NotI unique site

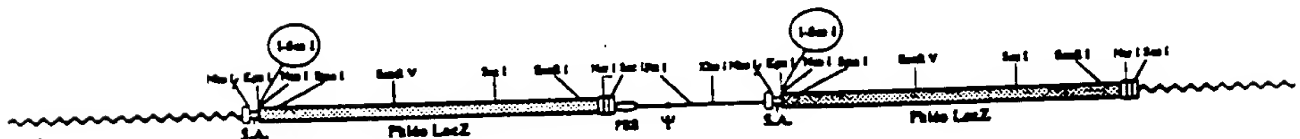
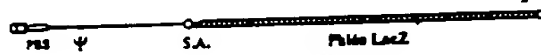
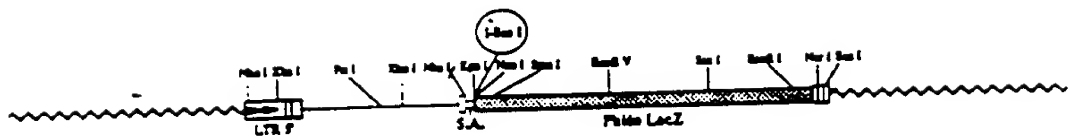
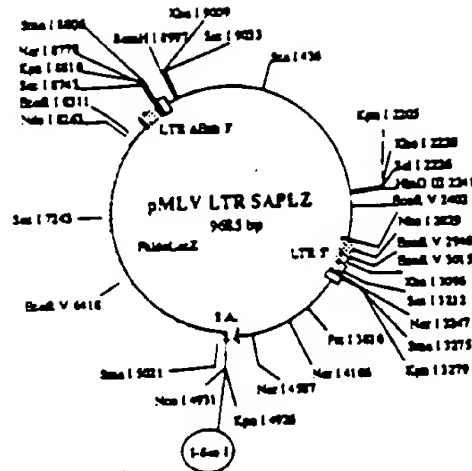
0002210" 26926460

FIGURE 12



Construction : pD 123 , from J. D. Boeke
with insertion of a linker [I-SceI - NotI] in BamHI

002270 26926460



0049697 012700

002270" 6926450

FIGURE 14

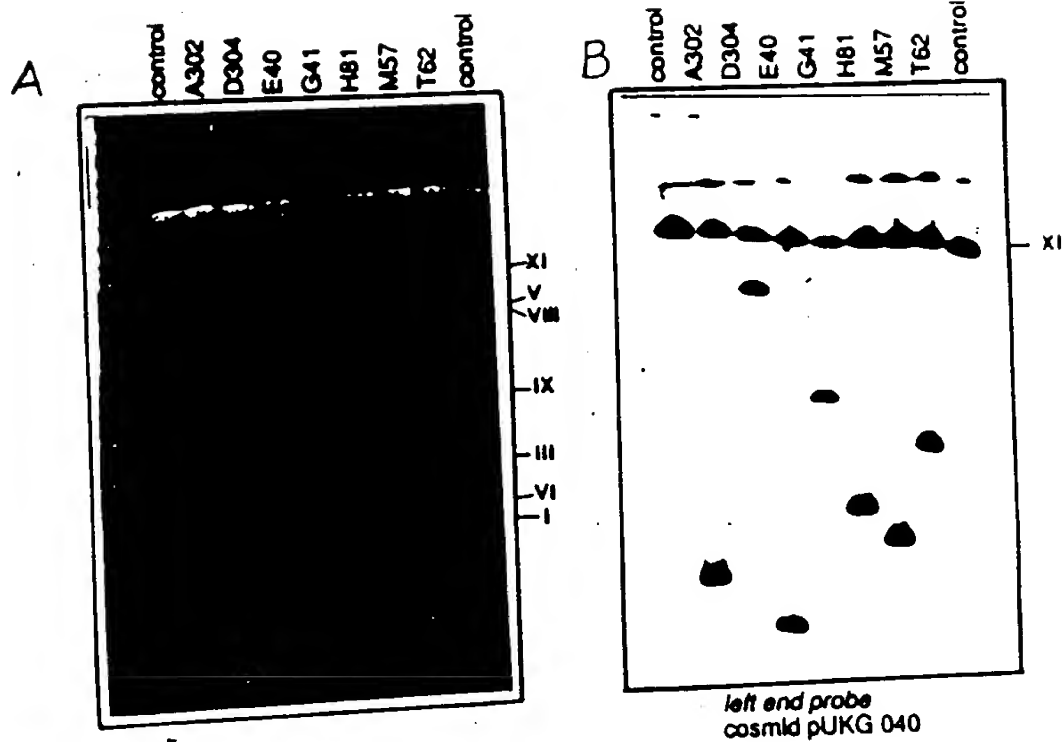
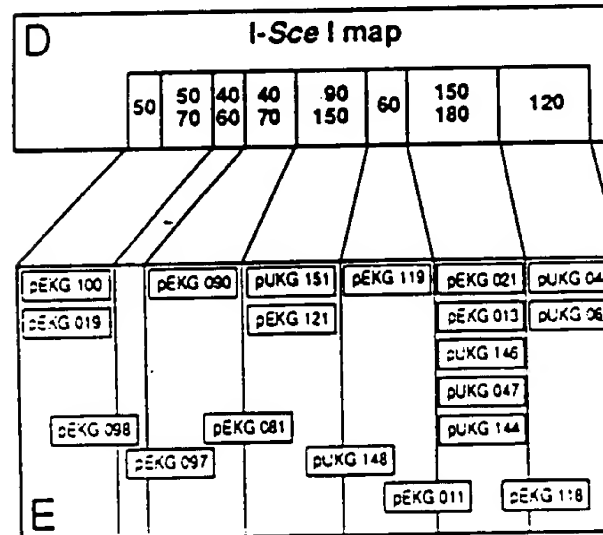
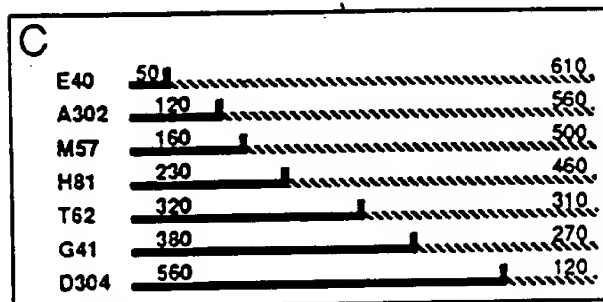
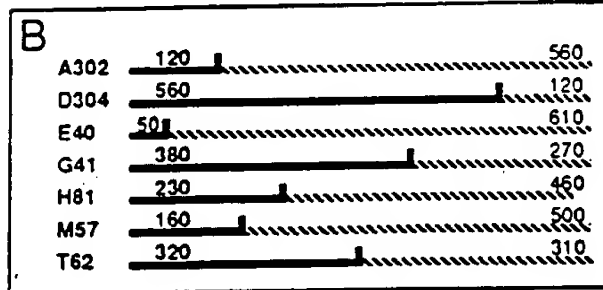
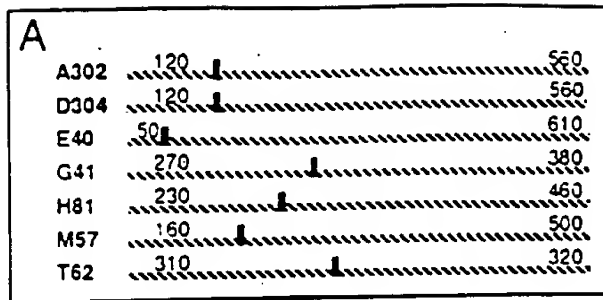


FIGURE 15



00/270/6326450

FIGURE 16

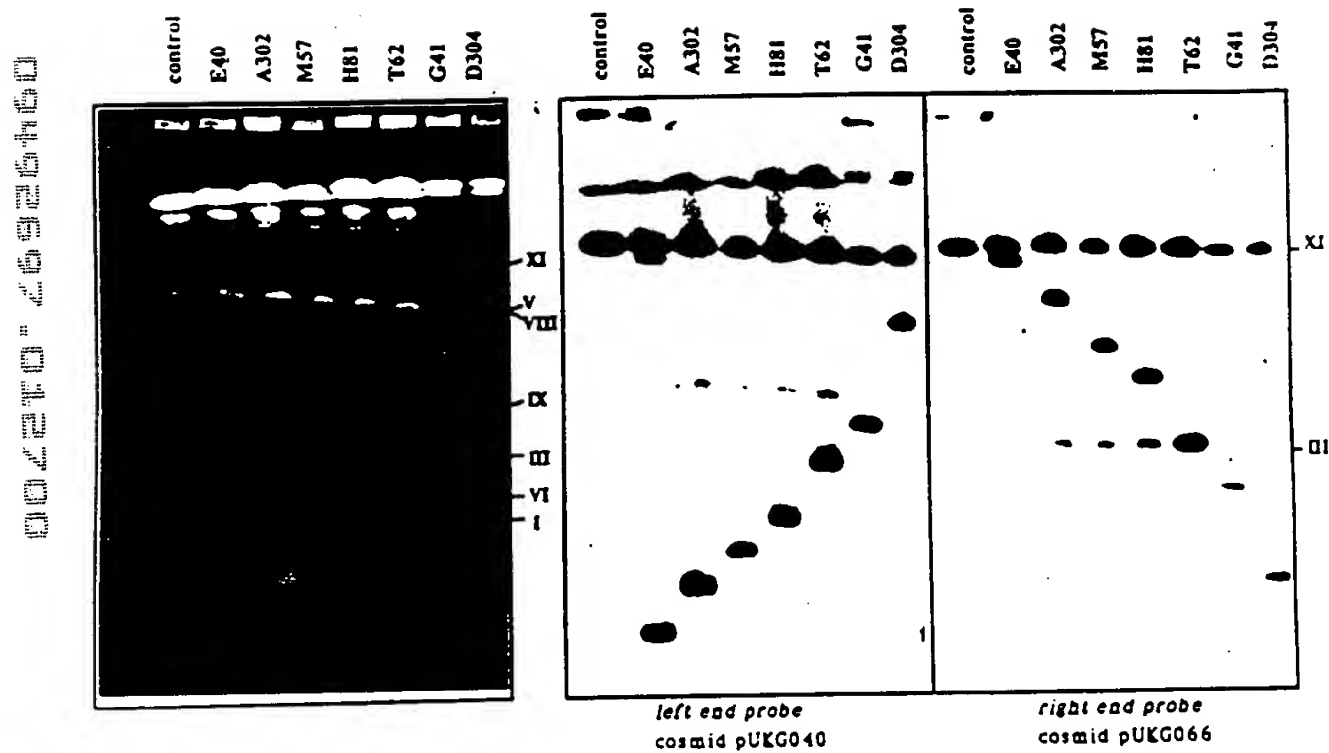
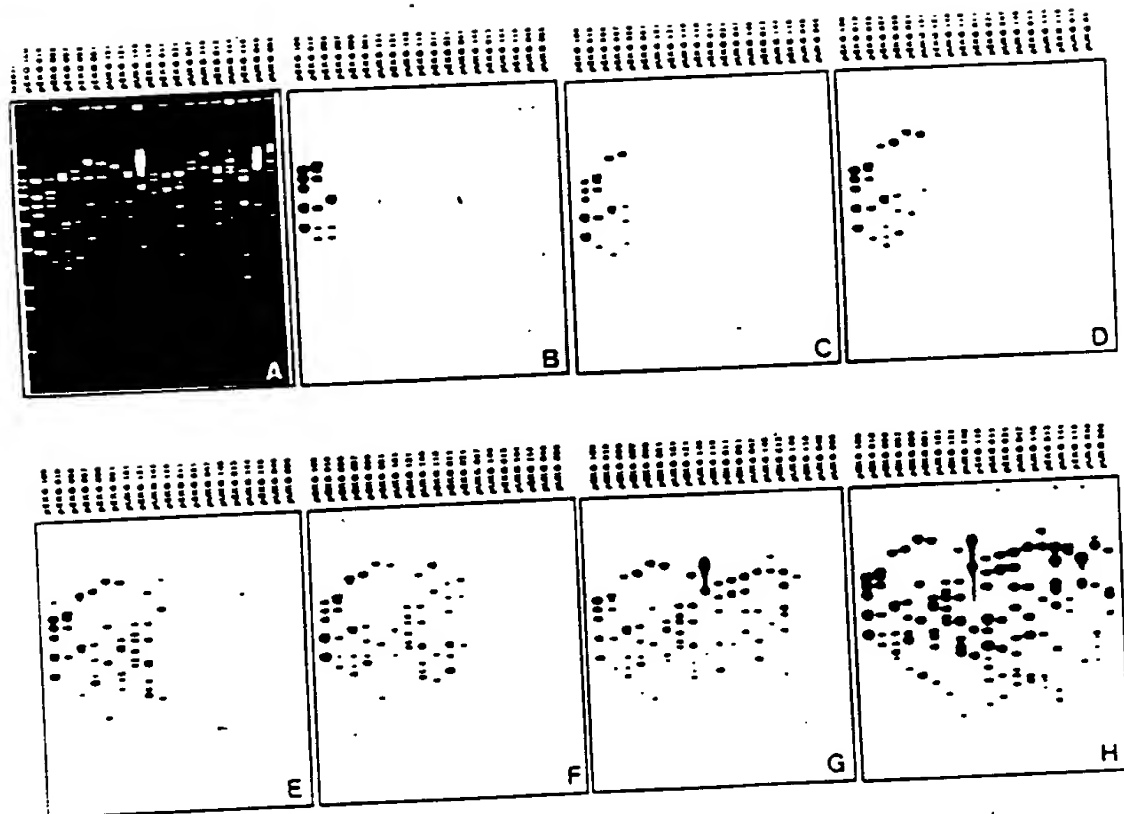


FIGURE 17

002270-6926460



0019657 012700

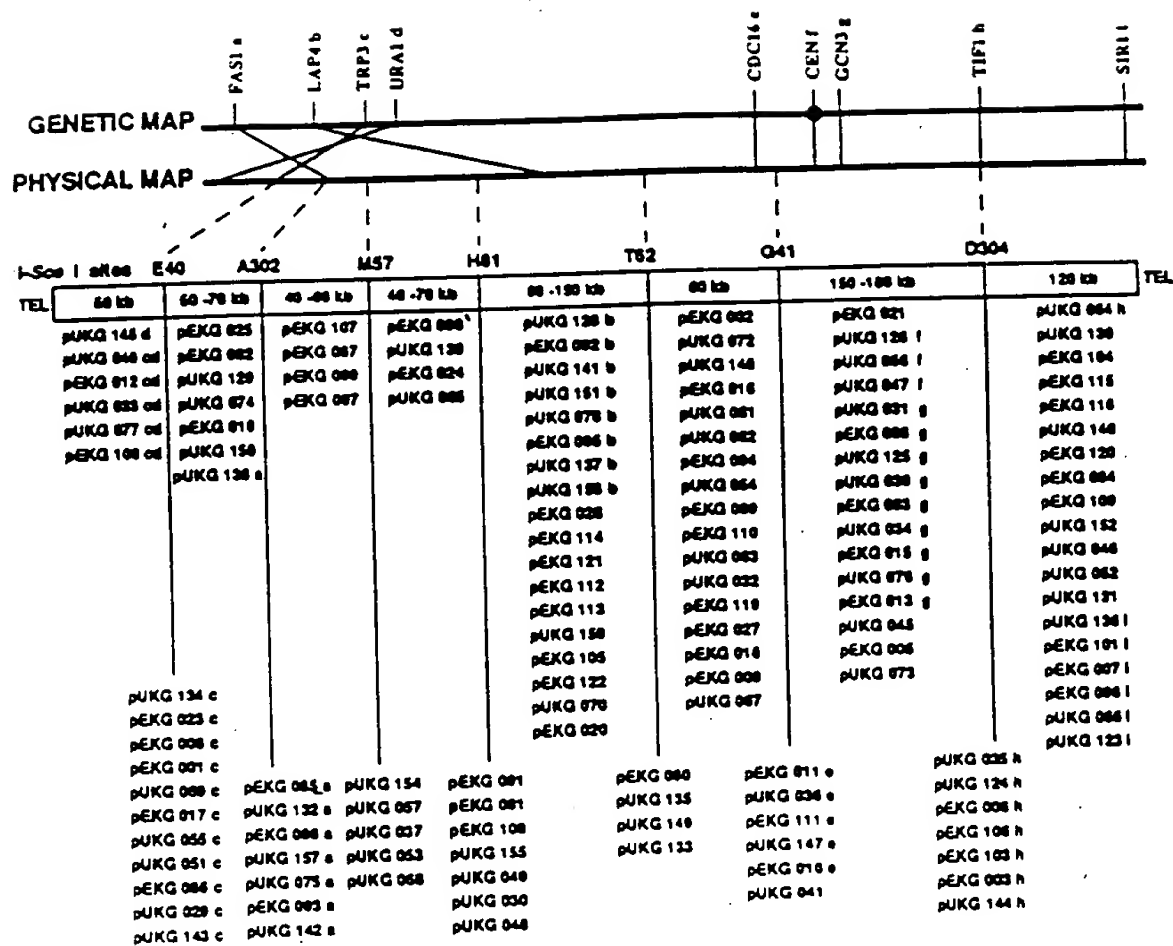
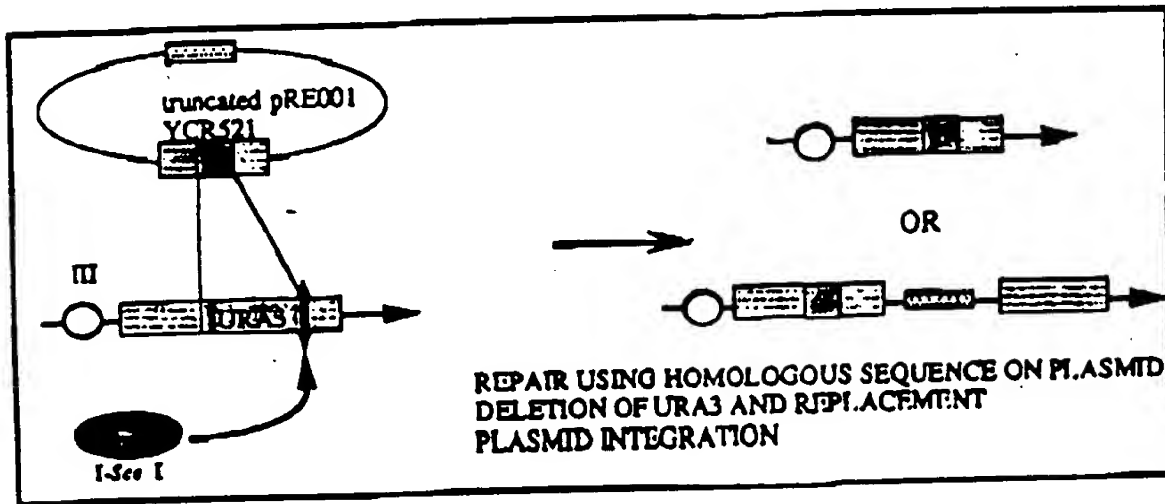
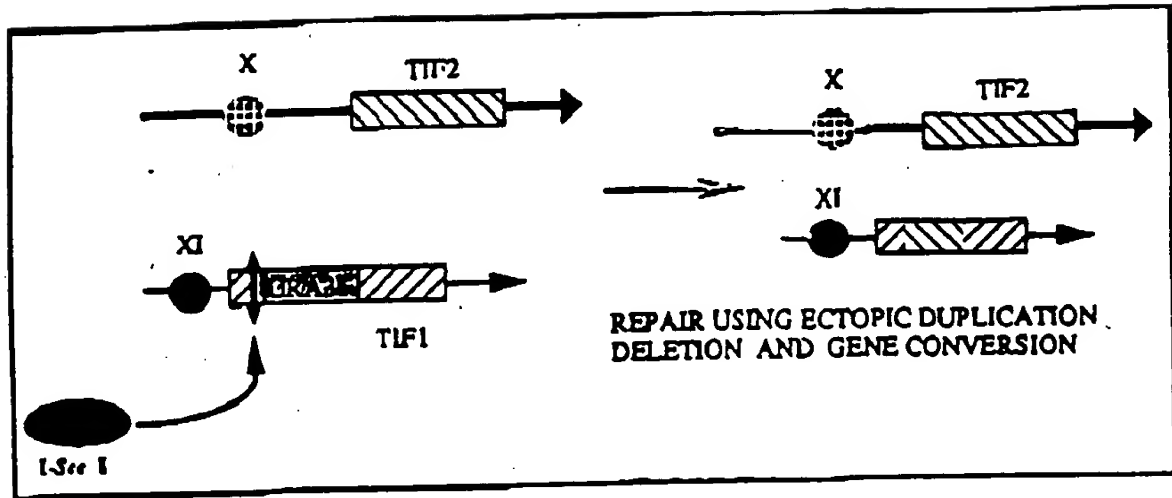
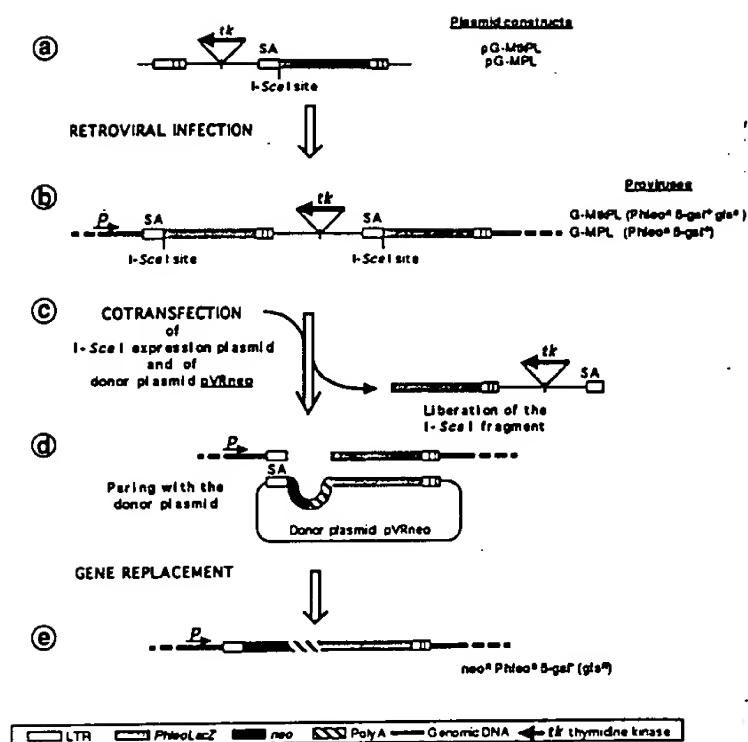


FIGURE 19



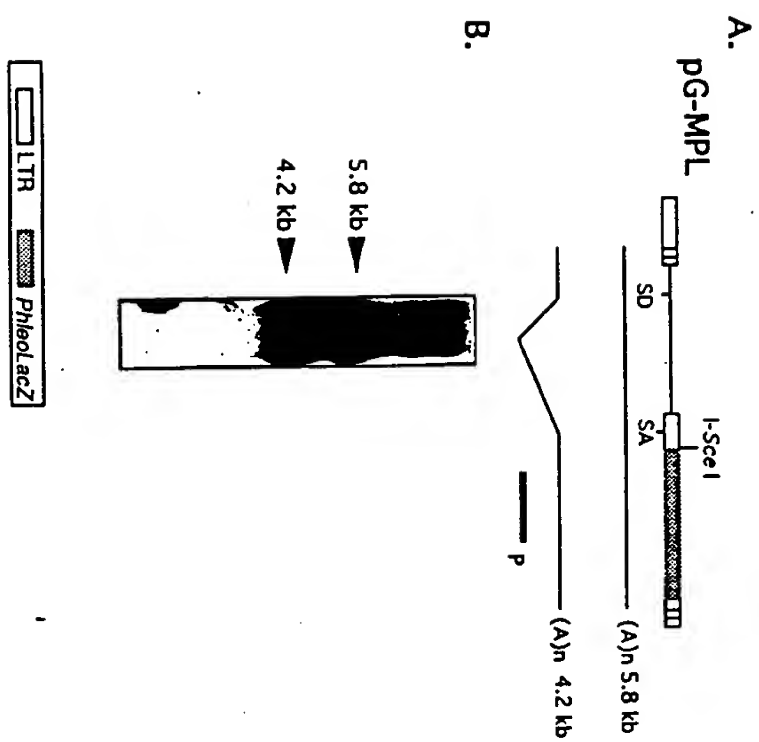
002270-26926460

Figure 20



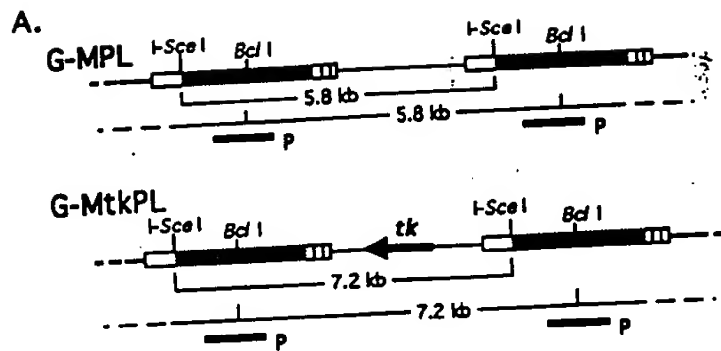
00492697-013700

Figure 21

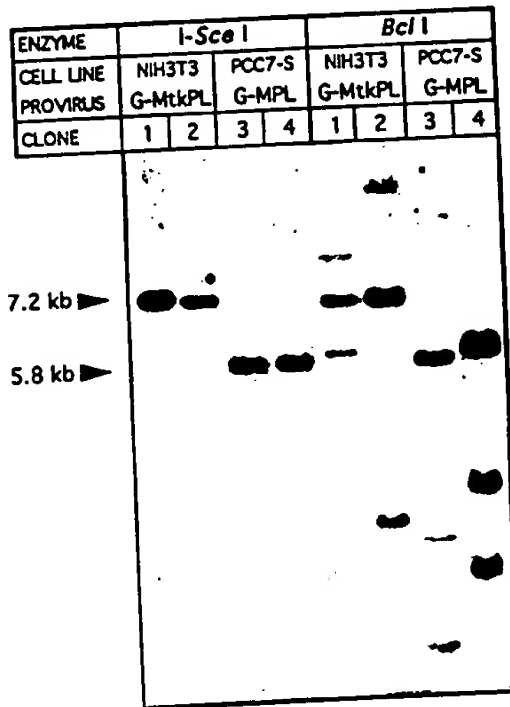


094492697 012700

Figure 22



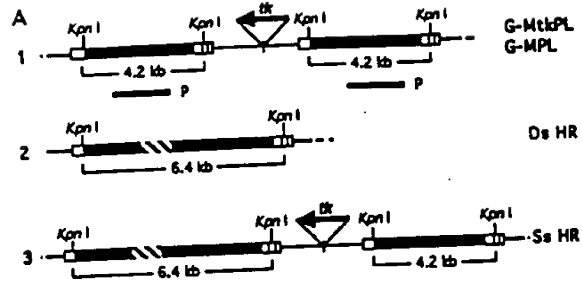
B.



LTR
 PhleoLacZ
 tk thymidine kinase

00492697-01200

Figure 23



B

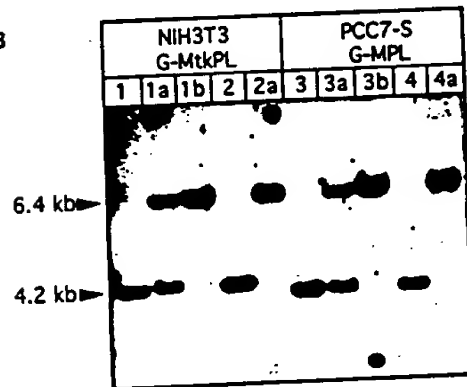
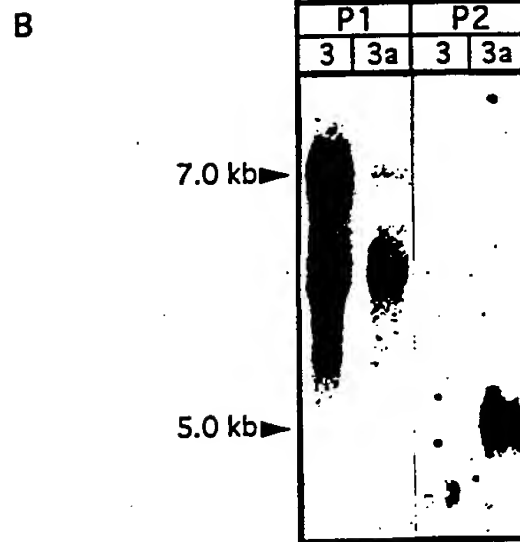
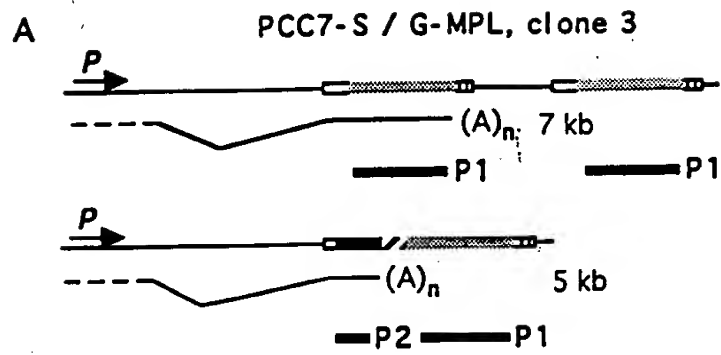


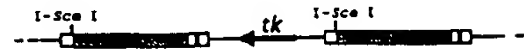
Figure 24



002270-46926460

Figure 25

a. Chromosomal DNA containing provirus

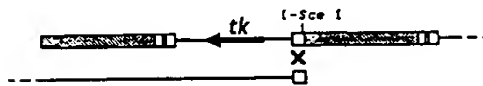


Phenotypes
[Phleo^R, Glu^R, 8-Gal⁺]

Transfection by
I-Sce I endonuclease
expression vector

b. Intra-chromosomal recombinations events

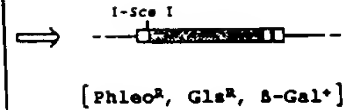
1. The left I-Sce I site is cut.
Pairing and recombination



2. The right I-Sce I site is cut.
Pairing and recombination



3. Both I-Sce I sites are cut.
Religation by end-joining



c. Inter-chromosomal recombination event

Both I-Sce I sites are cut. Gap repair using
intact chromosome sequences

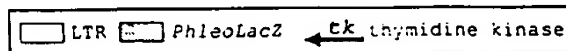
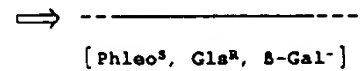
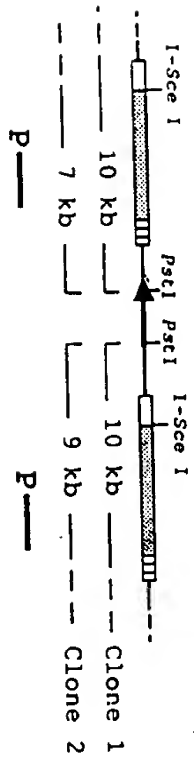
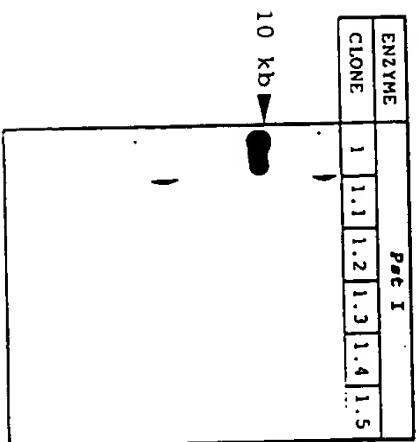


Figure 26

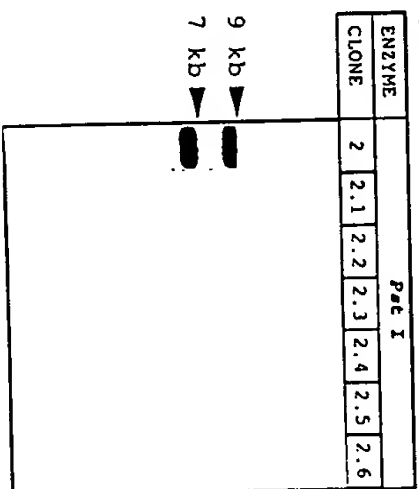
a. Parental DNA, G-MtkPL



b.



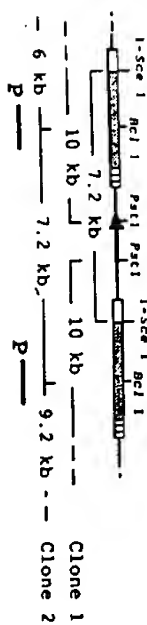
c.



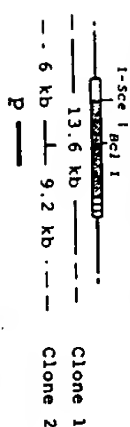
09492697 012700

Figure 27

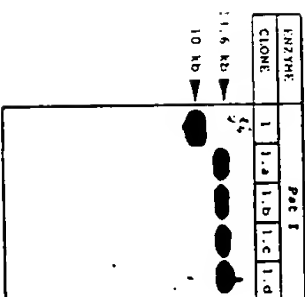
a. 1. Parental DNA, G-MtkPL



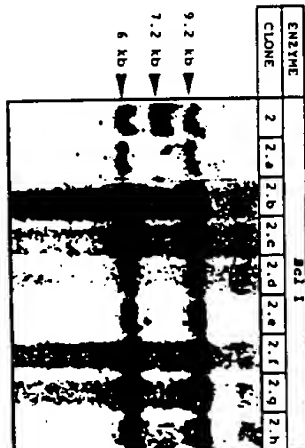
2. Intra-molecular recombination event



b.



c.



09492697 012700

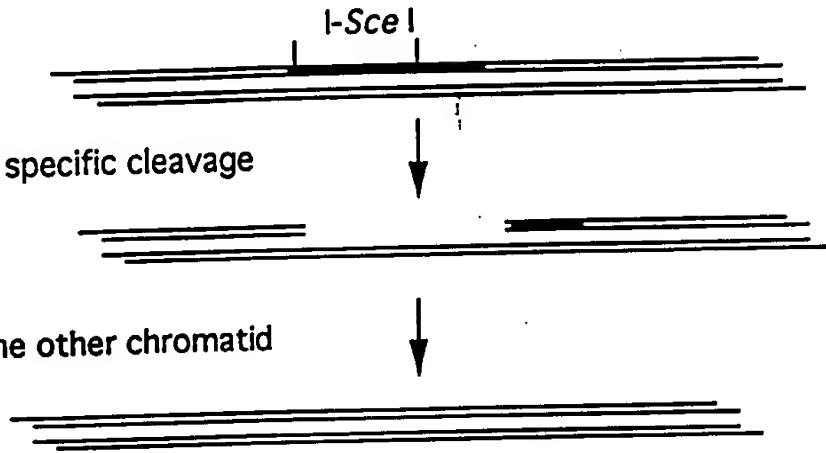
Figure 28

LOSS OF HETEROZYGOSITY

Integration of artificial site or presence of specific site

Expression of I-Sce I and specific cleavage

Repair of the DSB with the other chromatid

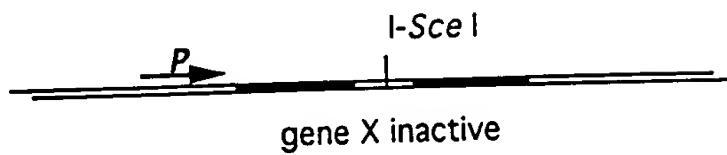


00492697-012700

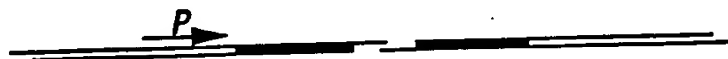
Figure 29

CONDITIONNAL ACTIVATION (Tandem repeat)

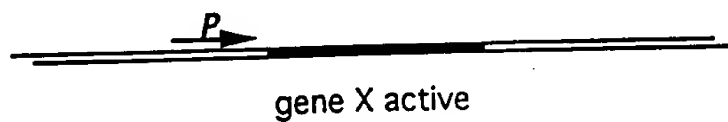
Integration of artificial site between tandem repeats



Expression of I-Sce I and specific cleavage

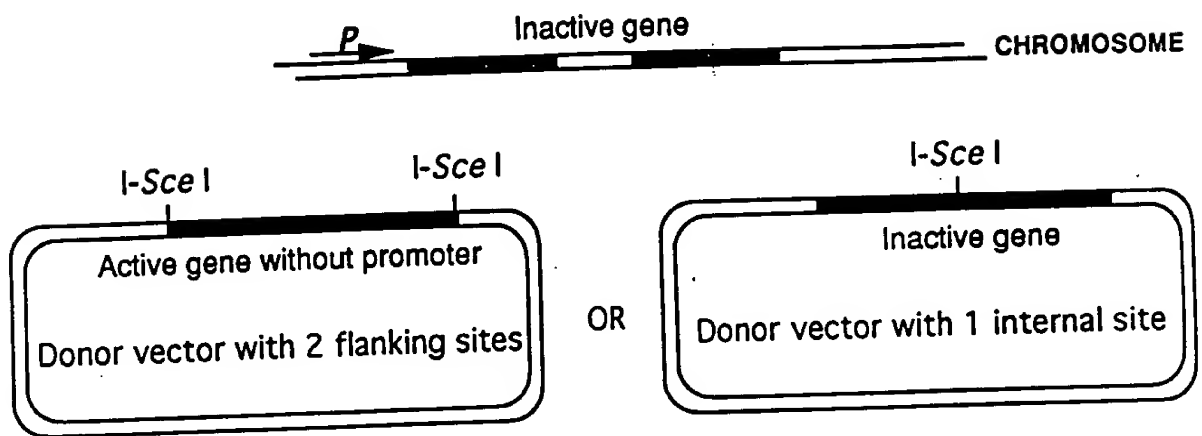


Repair of the DSB by single strand annealing

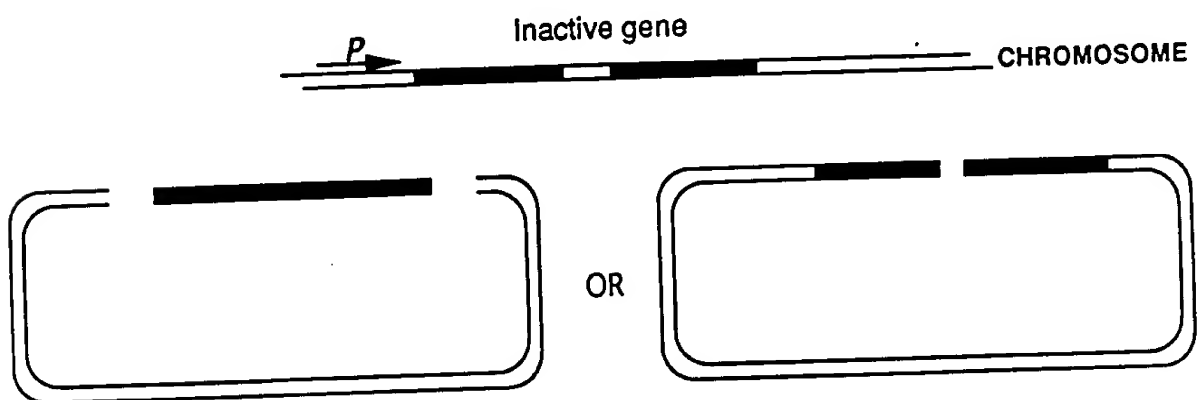


ONE STEP REARRANGEMENT

Integration of artificial site or presence of specific site



Expression of I-Sce I enzyme
and
specific cleavage of the donor plasmid



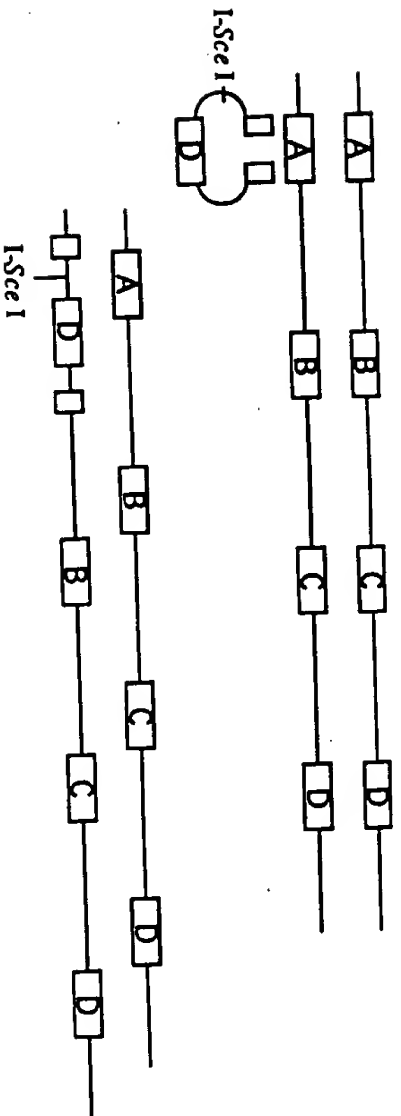
Recombination between the chromosome and plasmid



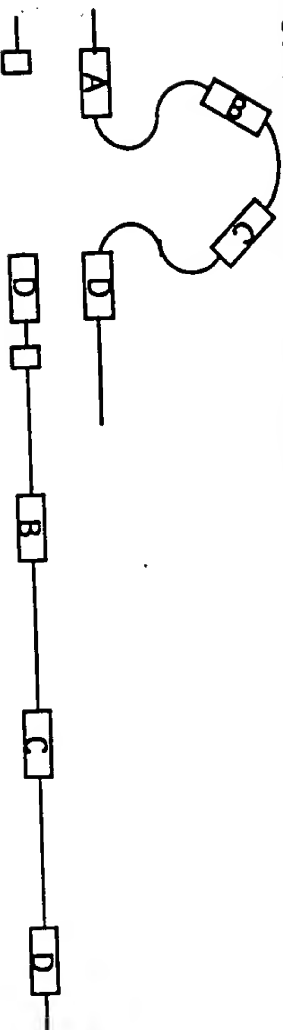
00492697-012700

DUPLICATION OF A LOCUS

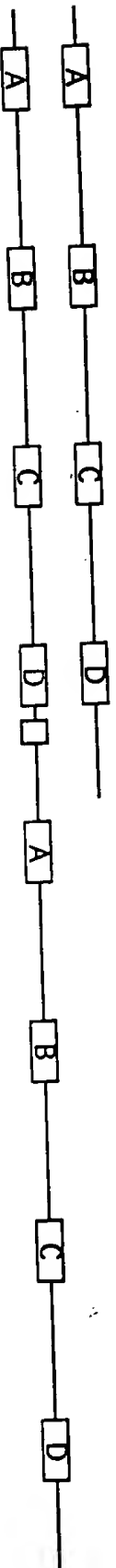
1 insertion of I-Sce I site by classical gene replacement



2 Specific cleavage by I-Sce I enzyme and repair of the break by homologous sequences



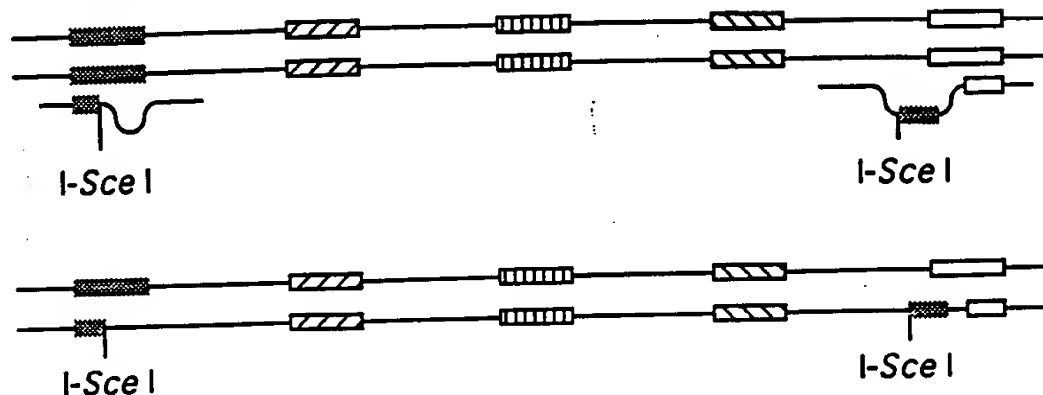
3 Duplication of the totality of the locus



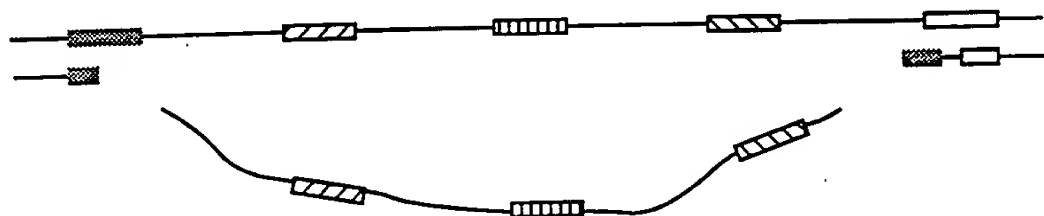
09492697 012700

DELETION OF A LOCUS

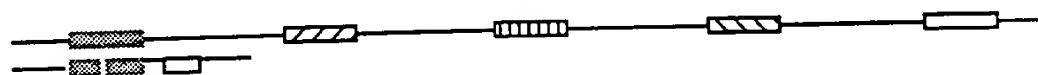
1 Insertion of two I-Sce I sites flanking the locus



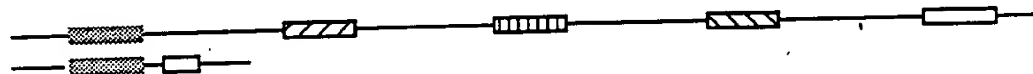
2 Expression of the enzyme and cleavage



3 Recombination between the two ends



4 deletion of the locus



00492637-012700

Figure 33

